Article

Genomic Prediction Accuracies for Growth and Carcass Traits in a Brangus Heifer Population

Sunday O. Peters [1,*] **, Kadir Kızılkaya** [2] **, Mahmut Sinecen** [3] **, Burcu Mestav** [4] **, Aranganoor K. Thiruvenkadan** [5] **and Milton G. Thomas** [6]

[1] Department of Animal Science, Berry College, Mount Berry, GA 30149, USA
[2] Department of Animal Science, Faculty of Agriculture, Aydin Adnan Menderes University, Aydin 09100, Turkey
[3] Department of Computer Engineering, Faculty of Engineering, Aydin Adnan Menderes University, Aydin 09100, Turkey
[4] Department of Statistics, Faculty of Arts and Sciences, Çanakkale Onsekiz Mart University, Terzioğlu Campus, Çanakkale 17100, Turkey
[5] Department of Animal Genetics and Breeding, Veterinary College and Research Institute, Tamil Nadu Veterinary and Animal Sciences University, Salem 637002, Tamil Nadu, India
[6] Texas A&M AgriLife Research, Beeville, TX 78102, USA
[*] Correspondence: speters@berry.edu; Tel.: +1-706-368-6919

Simple Summary: The genomic estimated breeding value (GEBV) using data from Brangus heifers were obtained from genomic selection (GS) methods associating the single nucleotide polymorphisms (SNP) marker genotypes with phenotypic data for economically important growth (birth, weaning, and yearling weights) and carcass (depth of rib fat, and percent intramuscular fat and longissimus muscle area) traits using the linkage disequilibrium (LD) between SNP markers and quantitative trait loci (QTL) and/or the genomic relationship between animals. The heritability estimates were found similar across genomic best linear unbiased prediction (the GBLUP), and the Bayesian (BayesA, BayesB, BayesC and Lasso) GS methods for k-means and random cluster. The Bayesian methods resulted in underestimates of heritabilities and overestimates of accuracy of GEBV. However, the GBLUP method resulted in more reasonable estimates of heritabilities and accuracies of GEBV for growth and carcass traits of heifers from a composite population.

Abstract: The predictive abilities and accuracies of genomic best linear unbiased prediction (GBLUP) and the Bayesian (BayesA, BayesB, BayesC and Lasso) genomic selection (GS) methods for economically important growth (birth, weaning, and yearling weights) and carcass (depth of rib fat, apercent intramuscular fat and longissimus muscle area) traits were characterized by estimating the linkage disequilibrium (LD) structure in Brangus heifers using single nucleotide polymorphisms (SNP) markers. Sharp declines in LD were observed as distance among SNP markers increased. The application of the GBLUP and the Bayesian methods to obtain the GEBV for growth and carcass traits within k-means and random clusters showed that k-means and random clustering had quite similar heritability estimates, but the Bayesian methods resulted in the lower estimates of heritability between 0.06 and 0.21 for growth and carcass traits compared with those between 0.21 and 0.35 from the GBLUP methodologies. Although the prediction ability of the GBLUP and the Bayesian methods were quite similar for growth and carcass traits, the Bayesian methods overestimated the accuracies of GEBV because of the lower estimates of heritability of growth and carcass traits. However, GBLUP resulted in accuracy of GEBV for growth and carcass traits that parallels previous reports.

Keywords: accuracy; GBLUP; Bayesian methods; genomic prediction; k-means clustering; growth and carcass traits

1. Introduction

The availability of high-density SNP genotypes from high-throughput genotyping technologies [1–4] and the development of linear and nonlinear methods (such as the GBLUP, BayesA, BayesB, BayesC, and Bayesian Lasso) [1,5,6] have made genomic selection applicable for the economically important traits in animal and plant breeding [7–14]. Genomic selection methods associate SNP marker genotypes with phenotypic data for economically important traits to obtain the GEBV of animals based on the LD between SNP and QTL and/or genomic relationship among animals. The accuracy of GEBV, important for the genetic progress in GS, is influenced by many factors, including the level of LD between SNP and QTL, heritability of the trait, and the estimation methods of GEBV [15–17]. Habier et al. [18] reported that the accuracies of GEBV depend on LD among SNP and QTL, and on genomic relationships among animals in the training and validation datasets. Their findings indicated that the accuracy of GEBV of a selected animal decreased as the genomic relationship between selection animals (candidates) and training animals decreased. Saatchi et al. [19] also showed that if the genetic relationships between animals in training and animals in validation data were minimized as per the pedigree-base additive genetic relationships among animals in the k-means clustering procedure, accuracies of GEBV of animals in the validation data were less affected by their genomic relationships. Villumsen et al. [20] also studied the effect of heritability on the accuracy of GEBV in GS using simulated data and reported that the accuracy of GEBV increased about 17% as the heritability increased from 0.02 to 0.30 in the GS study. Clark et al. [21] compared the accuracy of GEBV from BLUP, the GBLUP, and the BayesB methods, finding that the accuracies of genomic prediction from GS methods depended on the significant effect of QTL on the trait, and that the small effect of QTL resulted in a non-significant difference between GBLUP and BayesB.

The objectives of this research were to characterize LD structure of Brangus heifers and to compare the predictive ability and accuracy of the GBLUP and the Bayesian methods for economically important growth (birth, weaning, and yearling weights) and carcass (depth of rib fat, percent of intramuscular fat, and longissimus muscle area) traits using BovineSNP50 Infinium BeadChip SNP markers (n = 54,001 SNP).

2. Materials and Methods

2.1. Phenotypes

Birth weight (BW), weaning weight (WW), and yearling weight (YW) were phenotypes for growth traits, and depth of rib fat (FAT), percent intramuscular fat (IMF), and longissimus muscle area (LMA) were phenotypes for carcass traits from yearling ultrasound evaluation. Phenotypes were collected from 738 Brangus heifers that were registered with International Brangus Breeders Association [9,22,23]. Year of birth (2005 to 2007), season of calving (spring or autumn), and age of dam were also obtained from the database of the International Brangus Breeders Association. The descriptive statistics of these growth and carcass traits are presented in Table 1.

Table 1. Descriptive statistics for growth and carcass traits in Brangus heifers.

Trait	Mean ± SE *	Minimum	Maximum
Birth weight (BW), kg	34.48 ± 0.19	18.03	50.94
Weaning weight (WW), kg	377.98 ± 1.81	201.37	549.56
Yearling weight (YW), kg	540.67 ± 3.45	225.51	769.67
Depth of rib fat (FAT), cm	0.57 ± 0.01	0.02	1.40
Intramuscular fat (IMF), %	4.81 ± 0.04	2.02	9.77
Longissimus muscle area (LMA), cm^2	62.31 ± 0.41	27.22	91.19

* SE: Standard error of mean.

2.2. SNP Marker Genotypes

BovineSNP50 Infinium BeadChips for 54,001 SNP markers were used to genotype each heifer [2]. Genotypes of SNP markers were determined in the A/B allele format and

coded as 0, 1, or 2, based on the number of B alleles at each locus. With this SNP marker information and using the snpReady package in R-program [24], three filters were applied for quality control in the following sequence: (a) Animals with > 50% missing data were removed; (b) SNP markers with > 5% missing data or < 95% call rate were removed; and (c) SNP markers with < 10% minor allele frequency were removed. After executing imputation for missing SNP markers, the complete SNP genotype data included 35,351 SNP markers from 738 animals. On each chromosome, the distribution of the number of SNP markers within a 1 Mb window was determined by using the rMVP package in the R-program [25].

2.3. Linkage Disequilibrium

The success of GS and genome-wide association studies (GWAS) are dependent on LD, which is a non-random association among SNP markers. LD is measured using the square of correlation (r^2) between SNP markers and ranges between 0 and 1. Linkage disequilibrium is expressed as

$$r_{ij}^2 = \frac{(p_{AB} - p_A p_B)^2}{(p_A p_a)(p_B p_b)} \tag{1}$$

where p_{AB}, $p_A = 1 - p_a$ and $p_B = 1 - p_b$ are the observed frequencies for haplotype AB and alleles A and B at locus i and j, respectively. The estimates of LD for pairwise combinations of all SNP markers were obtained from the *pairwise LD* function of the Synbreed package in the R program [26,27].

2.4. Genomic Selection

2.4.1. Genomic Best Linear Unbiased Prediction (GBLUP)

The model for GEBV was:

$$y = Xb + Zg + e \tag{2}$$

where y was a vector of BW, WW, YW, LMA, IMF, or FAT; X was a design matrix allocating BW, WW, YW, LMA, IMF or FAT to the fixed effects of overall mean, contemporary groups and dam age; Z was a design matrix allocating BW, WW, YW, LMA, IMF, or FAT to additive genetic effects of animals; b was a vector of fixed effects of overall mean, contemporary groups, and dam age; and g was a vector of additive genomic breeding values for animals following a multivariate normal distribution $g \sim N\left(0, G\sigma_g^2\right)$ with genomic relationship matrix (G) and the additive genetic variance (σ_g^2) among animals. e was a vector of residuals following a multivariate normal distribution $e \sim N\left(0, I\sigma_e^2\right)$ with the residual variance (σ_e^2).

The G matrix indicating the realized relatedness among animals was calculated as

$$G = \frac{WW^T}{2\sum_{i=1}^{k} p_i(1 - p_i)} \tag{3}$$

where $W = M - P$, M was the ($n \times k$) matrix of SNP markers for the $n = 738$ animals with the $k = 35,351$ SNP markers; P was the ($n \times k$) matrix of the allele frequencies multiplied by 2; p_i was the allele frequency of SNP marker i; and the sum was, overall, loci [18,28].

The GBLUP used for the GEBV of animals was equivalent to solving the mixed model equations:

$$\begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + G^{-1}\frac{\sigma_e^2}{\sigma_g^2} \end{bmatrix} \begin{bmatrix} b \\ g \end{bmatrix} = \begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} \tag{4}$$

where σ_g^2 and σ_e^2 were the additive genetic and residual variances and G^{-1} was the inverse of the G matrix. Therefore, the heritability of the trait was defined as $h^2 = \sigma_g^2 / \left(\sigma_g^2 + \sigma_e^2\right)$.

The BGLR package (https://cran.r-project.org/web/packages/BGLR/index.html (accessed on 10 March 2022)) in the R program [6,26] was used to solve the mixed model

equations in Equation (4) for **b** and **g** by estimating the additive genetic and residual variances ($\hat{\sigma}_g^2$ and $\hat{\sigma}_e^2$). The estimate of heritability was then calculated as $\hat{h}^2 = \hat{\sigma}_g^2 / \left(\hat{\sigma}_g^2 + \hat{\sigma}_e^2 \right)$.

2.4.2. The Bayesian BayesA, BayesB, BayesC and Lasso Methods

The Bayesian (BayesA, BayesB, BayesC, and Lasso) methods were applied to estimate the SNP effects for genomic prediction using cross-validation datasets of BW, WW, YW, LMA, IMF, and FAT. The cross-validation data of BW, WW, YW, LMA, IMF, and FAT were modeled as a function of the individual SNP effects:

$$y = Xb + Mm + e \tag{5}$$

where **y** was a vector of BW, WW, YW, LMA, IMF, or FAT; **X** was a design matrix allocating BW, WW, YW, LMA, IMF, or FAT to the corresponding fixed effects of overall mean, contemporary groups, and dam age; **M** was a $n \times k$ matrix of SNP (0, 1 or 2); **b** was a vector of fixed effects of overall mean and contemporary groups and dam age; and **m** was a $k \times 1$ vector of SNP effects assumed a priori to follow a multivariate normal distribution **m** $\sim N(\mathbf{0}, \mathbf{\Omega})$ with $\mathbf{\Omega} = diag\left(\sigma_{m_1}^2, \sigma_{m_2}^2, \cdots, \sigma_{m_k}^2 \right)$ the diagonal matrix and $\sigma_{m_i}^2$ the variance of SNP i. The prior distribution of SNP effect m_i depended on the SNP variance $\sigma_{m_i}^2$ and the prior probability π that SNP i had zero effect:

$$m_i \bigg| \pi, \sigma_{m_i}^2 \begin{cases} = 0 & \text{with probability } \pi \\ \sim N\left(0, \sigma_{m_i}^2\right) & \text{with probability } (1 - \pi). \end{cases} \tag{6}$$

where the parameter of π was defined between 0 and 1 [5]. The specifications for π and the SNP variance $\sigma_{m_i}^2$ determined the methods of BayesA, BayesB, and BayesC. In BayesA and the BayesB methods, the SNP variance $\sigma_{m_i}^2$ denoted the ith SNP variance, which had a scaled inverse chi-square distribution ($\chi^{-2}(\nu, S)$) with degrees of freedom ν and scale S parameters. These specifications result in a univariate Student's t distribution $t(0, \nu, S)$ for the marginal distribution of the SNP effect $m_i | \nu, S$ with the probability of the parameter of $(1 - \pi)$ [5,6]. In BayesC, with the SNP variance $\sigma_{m_i}^2 = \sigma_m^2$, prior distributions of the SNP effects had a common variance distributed with $\chi^{-2}(\nu, S)$. Therefore, these specifications resulted in a mixture of multivariate Student's t distributions $t(\mathbf{0}, \nu, \mathbf{I}S)$ for the marginal distribution of the SNP effect $m_i | \nu, S$ with the probability parameter of $(1 - \pi)$ [5,6]. In the BayesA method, the value of zero was assigned for the parameter of π, resulting in all k SNP in the model. However, in the BayesB and the BayesC methods, the fixed value of 0.95 was assigned for the parameter of π, resulting in 5% of k SNP markers with none-null variances in the model. In the Bayesian Lasso (BL), all k SNP ($\pi = 0$) were in the model, as in the BayesA method, and each SNP marker variance $\sigma_{m_i}^2$ had a Laplace distribution $Exp\left(\frac{\lambda^2}{2}\right)$ with λ parameter, which had a conjugate prior distribution of Gamma. These specifications result in a Double Exponential (DE) distribution for the marginal distribution of SNP effect $m_i | \lambda^2$ with the probability the parameter of $(1 - \pi)$ [6,29]. The vector of **e** represented normally distributed residuals ($e \sim N\left(\mathbf{0}, \mathbf{I}\sigma_e^2\right)$) with the variance ($\sigma_e^2$), which has a $\chi^{-2}(\nu_e, S_e)$ with degrees of freedom ν_e and scale S_e parameters. The BGLR package (https://cran.r-project.org/web/packages/BGLR/index.html (accessed on 10 March 2022)) in the R program [6,26] was used to estimate SNP effects for BW, WW, YW, LMA, IMF, and FAT.

2.4.3. K-Means and Random Clustering

The animals for cross-validation were divided into 10-fold data sets by using the k-means clustering approach. K-means clustering maximizes genetic relatedness within each cross-validation set and minimizes it between cross-validation datasets based on

the genetic dissimilarity (*D*) matrix among animals [19], which was calculated from the pedigree numerator relationship (*A*) matrix [30]:

$$d_{ij} = 1 - \frac{a_{ij}}{\sqrt{a_{ii}a_{jj}}} = 1 - r_{ij} \tag{7}$$

where d_{ij} was a measure of genetic dissimilarity between individuals *i* and *j*, a_{ij} was the additive genetic relationship between individual *i*, and *j*, a_{ii} (a_{jj}) was the ith (j^{th}) diagonal element of the *A* matrix, which represented Wright's coefficient of relationship (r_{ij}) between individuals *i* and *j*. *GeneticsPed* package in the R-program [31] was used to create the pedigree numerator relationship (*A*) matrix, and the *factoextra* package in the R-program [32] implementing the Hartigan and Wong [33] algorithm was used for k-means clustering.

For random clustering, the animals were randomly divided into 10-fold datasets for cross-validation, and this procedure was replicated five times.

2.4.4. Accuracy of Genomic Prediction

The training process in the 10-fold cross-validation from k-means and random clustering approaches was performed by excluding one validation set to train on the remaining nine validation sets, and then the GEBV of animas in the omitted validation set were obtained. The predictive ability of the GBLUP and the Bayesian BayesA, BayesB, BayesC, and Lasso methods in the 10-fold datasets for cross-validation were determined using Pearson's correlation coefficient ($r_{y,\hat{y}}$) between the observed (*y*) and predicted (*ŷ*) phenotypic values for BW, WW, YW, LMA, IMF, and FAT.

The accuracy of *GEBV* represented the correlation ($r_{BV,GEBV}$) between the breeding values (*BV*) and *GEBV*. However, the BV of animals are unknown, and the accuracy of *GEBV* of animals for traits was calculated by pooling estimates from the 10-fold cross-validation strategy. The accuracy of the *GEBV* of animals for traits was estimated using Pearson's correlation coefficient ($r_{y,\hat{y}}$) weighted by the heritability (h^2) of the traits in the validation datasets [34]:

$$r_{BV,GEBV} = \frac{r_{y,\hat{y}}}{\sqrt{h^2}} \tag{8}$$

3. Results and Discussion

3.1. Distribution of SNP Markers and LD Analysis

We retained 35,351 SNP after filtering markers based on the quality-control criteria. The distribution and density plots of SNP markers per chromosome are presented in Figure 1A,B. The total length of the autosomal genome was 2509.0 Mb, with the shortest chromosome (i.e., 25) being 42.9 Mb in length and the longest chromosome (i.e., 1) being 158.2 Mb in length. The length of chromosome X was 148.6 Mb. As seen in Figure 1A, there was a decreasing trend in the number of SNP markers from chromosome 1 to chromosome X and the SNP coverage ranged between 620 (1.78%) on chromosome 25 and 2194 (6.31%) on chromosome 1. Chromosome 1 and 25 had the longest and the shortest chromosomes with 158.49 Mb and 42.91 Mb in a study of Sahiwal cattle [35] with 157.78 Mb and 42.21 Mb in Charolais, Limousine, and Blonde d'Aquitaine cattle [36], and with 158.03 Mb and 42.80 Mb in Vrindavani crossbred cattle in India [37]. Singh et al. [37] also reported that since the distribution of SNP was related with the length of chromosomes, chromosome 1 had the highest number of SNP (2798) and chromosome 25 had the least number of SNP (792). The largest distance between SNP markers was 3.26 Mb on chromosome 10, and the shortest distance was 0.01 kb on chromosome 15. The average distance between SNP markers was 57.24 kb. Lu et al. [38] reported that the total genome length for Angus, Charolais, and Crossbred beef cattle in Canada was between 2534.98 and 2535.30 Mb, with the shortest chromosome 25 being 42.72 Mb and the longest chromosome 1 being 158.09 Mb. The distribution of the number of SNP differed from 2026 to 2176 for the chromosome 1 and

from 580 to 607 for chromosome 28, and the overall average distance between two adjacent SNP markers was 70 kb.
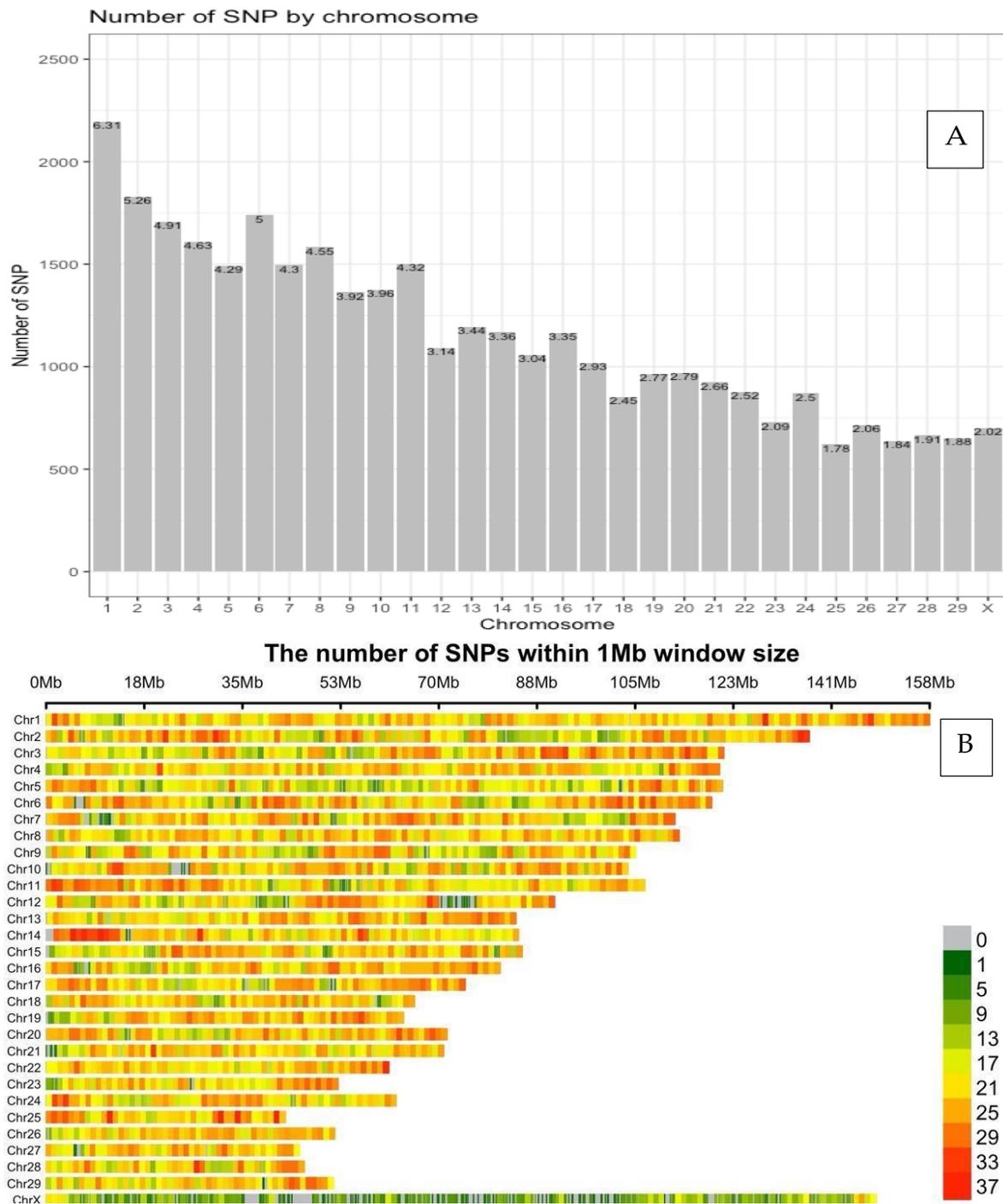

Figure 1. Distribution and density plot of SNP markers on each chromosome. (**A**): Number of SNP markers on each chromosome. (**B**): Density of SNP markers on each chromosome.

The density plot of SNP markers in Figure 1B showed the number of SNP markers within a 1 Mb window on each chromosome. The horizontal axis of the density plot of SNP markers indicates the length of chromosome (Mb). The different color shows SNP density from 0 to 37 SNP markers on each chromosome. The distribution of SNP markers on the autosomal chromosomes was not uniform and indicated a tendency of being clustered in some regions. The colors on the chromosomes showed the variation in the density of SNP

markers on each chromosome. The density of the SNP markers differed from 12.0 SNP/Mb on chromosome 12 to 15.1 SNP/Mb on chromosome 19. For the X chromosome, density of the SNP markers was 4.7 SNP/Mb. Chromosome 1 had a similar density pattern of SNP; however, chromosomes 11, 14, 24, and 25 had higher density of SNP at the beginning of the chromosomes compared to the rest of the chromosomes. The X chromosome was the second largest chromosome, but green and grey colors indicated very sparse densities of SNP markers. In addition, chromosome 6 had more SNP markers than chromosomes 3, 4, and 5, and was shorter than those chromosomes; therefore, the density of SNP markers on chromosome 6 (14.6 SNP/Mb) was higher than those on chromosomes 3 (14.1 SNP/Mb), 4 (13.3 SNP/Mb) and 5 (112.3 SNP/Mb).

Pairwise, LD between 35,351 SNP markers were assessed using the squared correlation (r^2) between SNP markers. The average LD (SD) and genetic distance (SD) were 0.125 (0.156) and 0.503 (0.285) Mb within an interval of 1 Mb pairs across all chromosomes. The overall average for LD and genetic distances were 0.022 (0.054) and 29.060 (24.209) Mb, respectively. The distribution of LD (r^2) against the genetic distance (Mb) given in Figure 2 indicated a sharp decline with increases of the genetic distance between SNP. The association between the degree of decay in LD with the distance between SNP markers indicated a clear decreasing exponential trend with an increasing genetic distance (Figure 2). Higher LD values were obtained for SNP markers located in close proximity. For the SNP markers less than 0.1 Mb apart, the mean LD (SD) was 0.195 (0.224), and 11.22% of SNP marker pairs had an LD higher than 0.5. For the genetic distance between pairs of SNP markers at ranges from 0 to 0.1, 0 to 0.2, and 0 to 0.5 Mb, 11.22, 8.73, and 5.67% of SNP marker pairs showed a higher LD than 0.5, respectively.
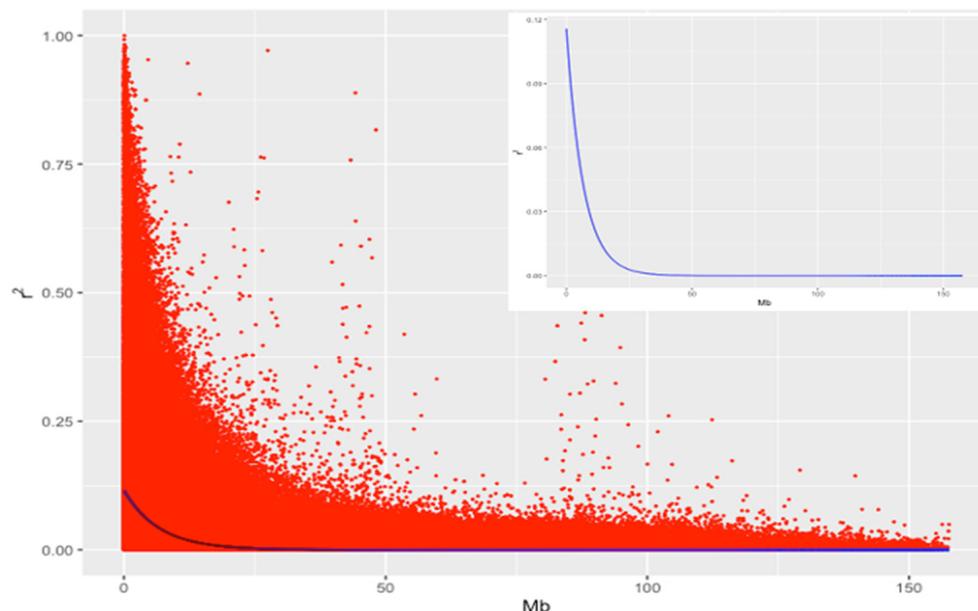
Figure 2. The linkage disequilibrium (r^2) among SNP markers plotted against the genetic distances (Mb) in Brangus heifers.

McKay et al. [39] and El Hou et al. [36] reported that most of the studies based on bovine SNP data have shown that the average LD was close to zero for distances between SNP greater than 500 kb. Lu et al. [38] reported rapidly decreasing LD from 0.29 to 0.23 to 0.19 in Angus, 0.22 to 0.16 to 0.12 in Charolais, and 0.21 to 0.15 to 0.11 in crossbred cattle for the distances from 0–30 kb to 30–70 kb, and then to 70–100 kb, respectively. El Hou et al. [36] also found that the average LD values between pairs of SNP markers ranged from 0.079 to 0.121 for Charolais, Limousine, and Blonde d'Aquitaine cattle, and the average LD changed from 0.5 to 0.1 at distances from smaller than 15 kb to greater than 120 kb.

Singh et al. [37] also calculated the average LD of 0.43 for the distance of less than 10 kb, and then it decreased to 0.21 for the distances of 25 to 50 kb for Vrindavani crossbred cattle.

3.2. Heritability Estimates from GBLUP, and the Bayesian (BayesA, BayesB, BayesC and Lasso) Methods in K-Means and Random Training Datasets

Heritability estimates of growth (BW, WW, and YW) and carcass (FAT, IMF, and LMA) traits from 10-fold k-means and random cluster training datasets were obtained from the GBLUP and the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods. In the analyses of growth traits, the estimates of heritabilities across 10-fold k-means (random) cluster training datasets ranged between 0.20 (0.20) to 0.26 (0.30) from the GBLUP and between 0.05 (0.04) to 0.16 (0.16) from the Bayesian methods for BW; between 0.19 (0.17) to 0.24 (0.24) from the GBLUP; between 0.02 (0.02) to 0.13 (0.13) from the Bayesian methods for WW; between 0.30 (0.31) to 0.36 (0.36) from the GBLUP; and between 0.09 (0.08) to 0.22 (0.29) from the Bayesian methods for YW. In the analyses of carcass traits, the estimates of heritabilities across 10-fold k-means (random) cluster training datasets ranged between 0.28 (0.28) to 0.34 (0.33) from the GBLUP and between 0.09 (0.07) to 0.19 (0.19) from the Bayesian methods for FAT; between 0.31 (0.30) to 0.37 (0.37) from the GBLUP, and between 0.15 (0.14) to 0.25 (0.26) from the Bayesian methods for IMF; and between 0.30 (0.32) to 0.38 (0.42) from the GBLUP, and between 0.11 (0.11) to 0.23 (0.29) from the Bayesian methods for LMA (Table 2). Overall mean (±standard deviation) estimates of heritabilities for the growth traits for 10-fold k-means (random) cluster training datasets in Figure 3 were 0.23 ± 0.02 (0.24 ± 0.02) from the GBLUP and 0.09 ± 0.03 (0.09 ± 0.02), 0.10 ± 0.02 (0.10 ± 0.03), 0.09 ± 0.01 (0.09 ± 0.01) and 0.15 ± 0.01 (0.15 ± 0.01) from BayesA, BayesB, BayesC, and BL in the Bayesian methods for BW; 0.22 ± 0.02 (0.21 ± 0.02) and 0.06 ± 0.03 (0.06 ± 0.03), 0.07 ± 0.03 (0.06 ± 0.02), 0.06 ± 0.01 (0.06 ± 0.01), 0.13 ± 0.01 (0.12 ± 0.01) for WW; 0.32 ± 0.02 (0.31 ± 0.03) and 0.16 ± 0.03 (0.16 ± 0.04), 0.16 ± 0.03 (0.15 ± 0.03), 0.15 ± 0.02 (0.15 ± 0.02), 0.17 ± 0.01 (0.17 ± 0.01) for YW, respectively.

Table 2. Mean [minimum, maximum] estimates of heritability of birth weight (BW), weaning weight (WW), and yearling weight (YW) for growth traits, rib fat (FAT), intramuscular fat (IMF), and longissimus muscle area (LMA) for carcass traits from 10-fold k-means and random cluster training datasets across replications, using Genomic Best Linear Unbiased Prediction (GBLUP) and Bayesian (BayesA, BayesB, BayesC, and Lasso (BL)) methods.

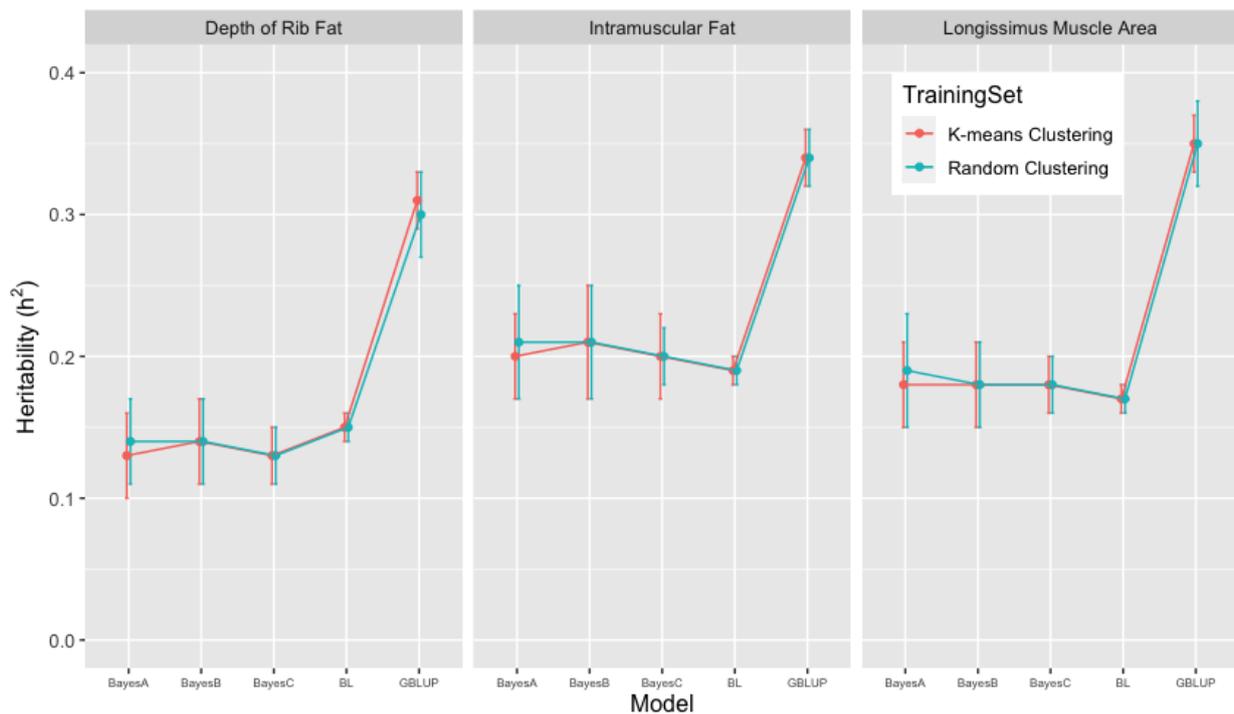| K-Means Cluster				
THE GBLUP	BayesA	BayesB	BayesC	BL
Growth Traits				
BW 0.23 [0.20, 0.26]	0.09 [0.05, 0.13]	0.10 [0.07, 0.13]	0.09 [0.06, 0.11]	0.15 [0.14, 0.16]
WW 0.22 [0.19, 0.24]	0.06 [0.02, 0.11]	0.07 [0.02, 0.11]	0.06 [0.05, 0.08]	0.13 [0.12, 0.13]
YW 0.32 [0.30, 0.36]	0.16 [0.11, 0.22]	0.16 [0.09, 0.20]	0.15 [0.11, 0.19]	0.17 [0.15, 0.18]
Carcass Traits				
FAT 0.31 [0.28, 0.34]	0.13 [0.09, 0.19]	0.14 [0.09, 0.17]	0.13 [0.10, 0.17]	0.15 [0.14, 0.16]
IMF 0.34 [0.31, 0.37]	0.20 [0.16, 0.24]	0.21 [0.15, 0.25]	0.20 [0.16, 0.23]	0.19 [0.18, 0.20]
LMA 0.35 [0.30, 0.38]	0.18 [0.13, 0.23]	0.18 [0.11, 0.21]	0.18 [0.12, 0.21]	0.17 [0.15, 0.18]
Random Cluster				
GBLUP	BayesA	BayesB	BayesC	BL
Growth Traits				
BW 0.24 [0.20, 0.30]	0.09 [0.04, 0.15]	0.10 [0.04, 0.20]	0.09 [0.06, 0.13]	0.15 [0.14, 0.16]
WW 0.21 [0.17, 0.24]	0.06 [0.02, 0.17]	0.06 [0.02, 0.11]	0.06 [0.04, 0.09]	0.12 [0.11, 0.13]
YW 0.31 [0.26, 0.38]	0.16 [0.08, 0.29]	0.15 [0.08, 0.25]	0.15 [0.11, 0.21]	0.17 [0.15, 0.18]
Carcass Traits				
FAT 0.30 [0.23, 0.35]	0.14 [0.07, 0.23]	0.14 [0.06, 0.23]	0.13 [0.09, 0.18]	0.15 [0.14, 0.16]
IMF 0.34 [0.29, 0.39]	0.21 [0.13, 0.28]	0.21 [0.13, 0.31]	0.20 [0.15, 0.25]	0.19 [0.17, 0.20]
LMA 0.35 [0.29, 0.42]	0.19 [0.11, 0.29]	0.18 [0.11, 0.28]	0.18 [0.13, 0.21]	0.17 [0.15, 0.18]
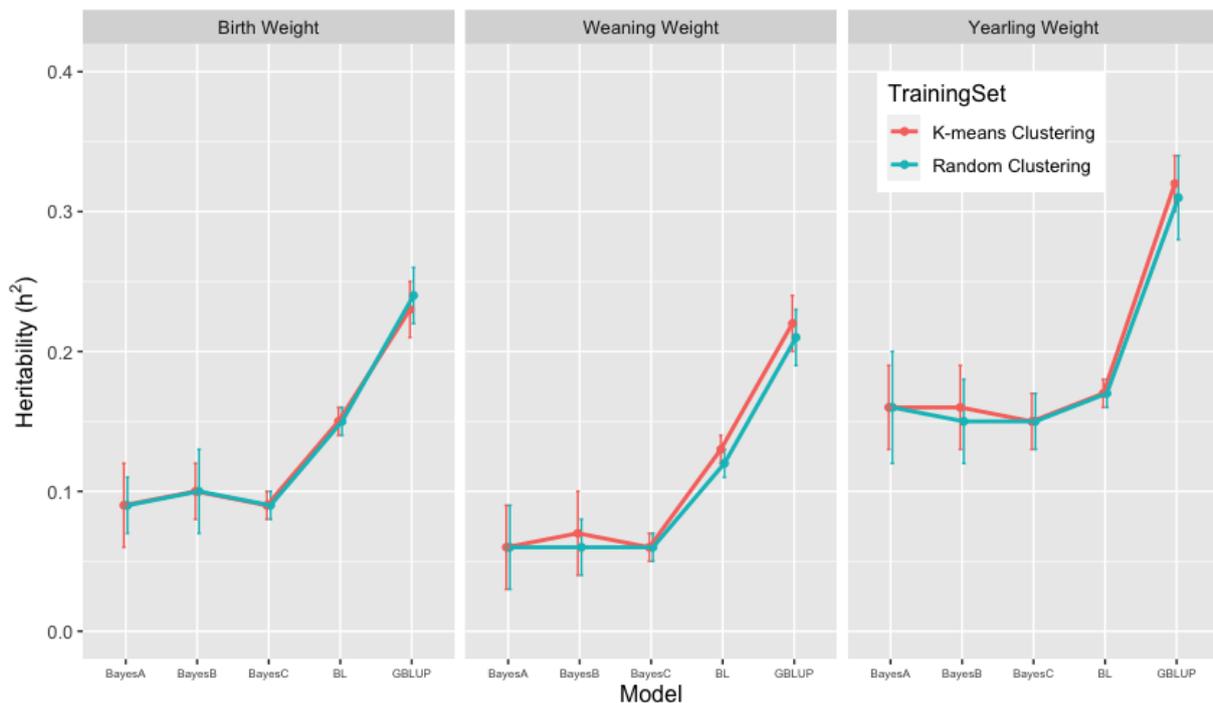
Figure 3. Mean estimates of heritability of birth weight (BW), weaning weight (WW) and yearling weight (YW) for growth traits from 10-fold k-means and random cluster training datasets using Genomic Best Linear Unbiased Prediction (GBLUP) and the Bayesian (BayesA, BayesB, BayesC, and Lasso (BL)) methods.

For the carcass traits, overall mean (±standard deviation) estimates of heritabilities in Figure 4 were 0.31 ± 0.02 (0.30 ± 0.03) from GBLUP and 0.13 ± 0.03 (0.13 ± 0.03), 0.14 ± 0.03 (0.14 ± 0.03), 0.13 ± 0.02 (0.13 ± 0.02) and 0.15 ± 0.01 (0.15 ± 0.01) from BayesA, BayesB, BayesC, and BL in the Bayesian methods for FAT; 0.34 ± 0.02 (0.34 ± 0.02) and 0.20 ± 0.03 (0.21 ± 0.04), 0.21 ± 0.04 (0.21 ± 0.04), 0.20 ± 0.03 (0.20 ± 0.02) and (0.19 ± 0.01) (0.19 ± 0.01) for IMF; 0.35 ± 0.02 (0.35 ± 0.03) and 0.18 ± 0.03 (0.18 ± 0.04), 0.18 ± 0.03 (0.18 ± 0.03), 0.18 ± 0.02 (0.18 ± 0.02) and 0.17 ± 0.01 (0.17 ± 0.01) for LMA.

As presented in Figures 3 and 4, 10-fold k-means and random cluster training datasets resulted in very similar heritability (h^2) estimates for growth and carcass traits. The comparison of methods suggested that the GBLUP methodology yielded almost double the heritability (h^2) estimates than the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods for growth and carcass traits within 10-fold k-means and random cluster training datasets. Within the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods, the BL method resulted in higher estimates of heritability (h^2) than BayesA, BayesB, and the BayesC methods for growth traits; however, heritability (h^2) estimates for carcass traits were similar across the Bayesian methods. Peters et al. [9] reported the pedigree and genome-based estimates of heritabilities for growth and carcass traits by conducting GWAS analyses using the BayesC method and the SNP markers for the Brangus cattle of this study.

Pedigree-based estimates of heritabilities were similar with those from GBLUP for growth traits, but were higher than those from the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods for growth and carcass traits. Genome-based (BayesC) estimates of heritabilies for growth and carcass traits were lower than those from GBLUP, but were similar with those from the other the Bayesian methods. The heritability (h^2) estimates from THE GBLUP for growth and carcass traits were in the range of heritability (h^2) estimates reported in the literature [9,40–44], and they suggested that GBLUP resulted in more reasonable heritability estimates than the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods in the analyses using smaller subsets of the data [42].

Figure 4. Means of heritability (h^2) estimates of depth of rib fat (FAT), intramuscular fat (IMF), and longissimus muscle area (LMA) for carcass traits from 10-fold k-means and random cluster training datasets using Genomic Best Linear Unbiased Prediction (GBLUP) and the Bayesian (BayesA, BayesB, BayesC, and Lasso (BL)) methods.

### 3.3. Comparison of Genome-Wide Prediction Ability

The GBLUP and the Bayesian methods were used to analyze growth (BW, WW, and YW) and carcass (FAT, IMF, and LMA) traits. The means and standard deviations of Pearson correlations ($r_{y,\hat{y}}$) between actual and predicted phenotypes for growth (BW, WW and YW) traits in Figure 5 and those for carcass (FAT, IMF, and LMA) traits in Figure 6 indicate the performance of genomic prediction from the GBLUP and the Bayesian methods when applied in the same 10-fold k-means and random cluster datasets for training and validation, respectively. As seen in Figures 5 and 6, mean correlations for growth (BW, WW and YW) and carcass (FAT, IMF and LMA) traits from k-means and random cluster training data sets were similar across the GBLUP and the Bayesian methods. For growth (BW, WW and YW) traits, Figure 5 showed that BL method (on average, 0.813 ± 0.005, 0.827 ± 0.005, 0.811 ± 0.005 and 0.814 ± 0.005, 0.827 ± 0.004, 0.811 ± 0.004) resulted in higher Pearson's correlations than GBLUP (on average, 0.747 ± 0.015, 0.745 ± 0.020, 0.801 ± 0.014 and 0.749 ± 0.020, 0.733 ± 0.024, 0.800 ± 0.020), BayesA (on average, 0.737 ± 0.044, 0.728 ± 0.075, 0.802 ± 0.029 and 0.746 ± 0.039, 0.729 ± 0.074, 0.808 ± 0.039), BayesB (on average, 0.751 ± 0.025, 0.733 ± 0.065, 0.799 ± 0.027 and 0.749 ± 0.043, 0.722 ± 0.058, 0.798 ± 0.033), and BayesC (on average, 0.747 ± 0.017, 0.743 ± 0.025, 0.799 ± 0.015 and 0.749 ± 0.022, 0.732 ± 0.028, 0.798 ± 0.020) methods within the 10-fold k-means and random cluster training datasets. For carcass (FAT, IMF and LMA) traits, Figure 6 showed that GBLUP (on average 0.812 ± 0.018, 0.814 ± 0.016, 0.827 ± 0.014 and 0.810 ± 0.022, 0.816 ± 0.015, 0.826 ± 0.021), BayesA (on average, 0.814 ± 0.029, 0.816 ± 0.018, 0.829 ± 0.028 and 0.820 ± 0.031, 0.822 ± 0.028, 0.834 ± 0.031), BayesB (on average, 0.817 ± 0.022, 0.822 ± 0.024, 0.825 ± 0.025 and 0.813 ± 0.033, 0.822 ± 0.025, 0.826 ± 0.022), BayesC (on average, 0.812 ± 0.020, 0.817 ± 0.017, 0.825 ± 0.015 and 0.811 ± 0.021, 0.818 ± 0.016, 0.825 ± 0.015), and BL (on average, 0.827 ± 0.006, 0.810 ± 0.003, 0.816 ± 0.004 and 0.827 ± 0.005, 0.811 ± 0.004, 0.817 ± 0.004) methods produced similar correlations ranging from 0.810 to 0.834 within the 10-fold k-means and random cluster

training datasets; however, the correlations for carcass traits were higher than those for growth traits.
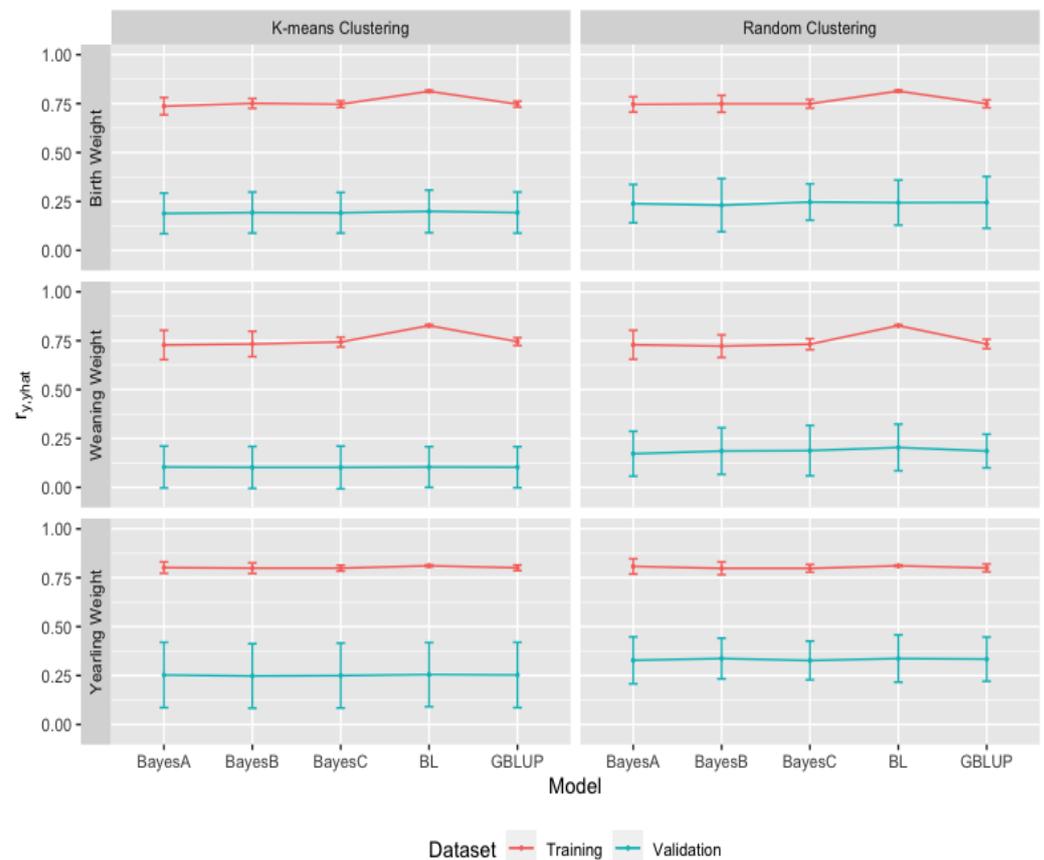
Figure 5. Predictive ability of Genomic Best Linear Unbiased Prediction (GBLUP) and the Bayesian (BayesA, BayesB, BayesC, and Lasso (BL)) methods for birth weight (BW), weaning weight (WW) and yearling weight (YW) for growth traits from 10-fold k-means and random cluster training datasets.

Predictive performance of the GBLUP and the Bayesian methods was explored by using the correlations from 10-fold k-means and random cluster cross-validation datasets. Figure 5 showed that GBLUP (on average, 0.193 ± 0.105, 0.103 ± 0.105, 0.253 ± 0.167 and 0.245 ± 0.132, 0.186 ± 0.086, 0.334 ± 0.113), BayesA (on average, 0.189 ± 0.104, 0.104 ± 0.107, 0.253 ± 0.167 and 0.239 ± 0.098, 0.172 ± 0.115, 0.328 ± 0.120), BayesB (on average, 0.193 ± 0.105, 0.102 ± 0.107, 0.248 ± 0.165 and 0.231 ± 0.136, 0.186 ± 0.119, 0.337 ± 0.104), BayesC (on average, 0.192 ± 0.104, 0.102 ± 0.109, 0.250 ± 0.166 and 0.247 ± 0.093, 0.188 ± 0.129, 0.327 ± 0.099), and BL (on average, 0.199 ± 0.109, 0.104 ± 0.104, 0.255 ± 0.164 and 0.244 ± 0.115, 0.204 ± 0.119, 0.337 ± 0.121) methods within the 10-fold k-means and random cluster cross-validation datasets resulted in similar correlations for growth (BW, WW and YW) traits. The ranges of the correlations within 10-fold k-means and random cluster cross-validations were from 0.189 ± 0.104 to 0.199 ± 0.109 and 0.231 ± 0.136 to 0.247 ± 0.093 for BW, 0.102 ± 0.107 to 0.104 ± 0.104 and 0.172 ± 0.115 to 0.204 ± 0.119 for WW, and 0.248 ± 0.165 to 0.255 ± 0.164 and 0.327 ± 0.099 to 0.337 ± 0.121 for YW. The ranges of correlations also indicated that the random cluster cross-validation resulted in a higher correlation than k-means cluster cross-validation, minimizing the genetic relationships among clusters. The trait of YW produced higher correlations than the traits of BW and WW within growth traits.
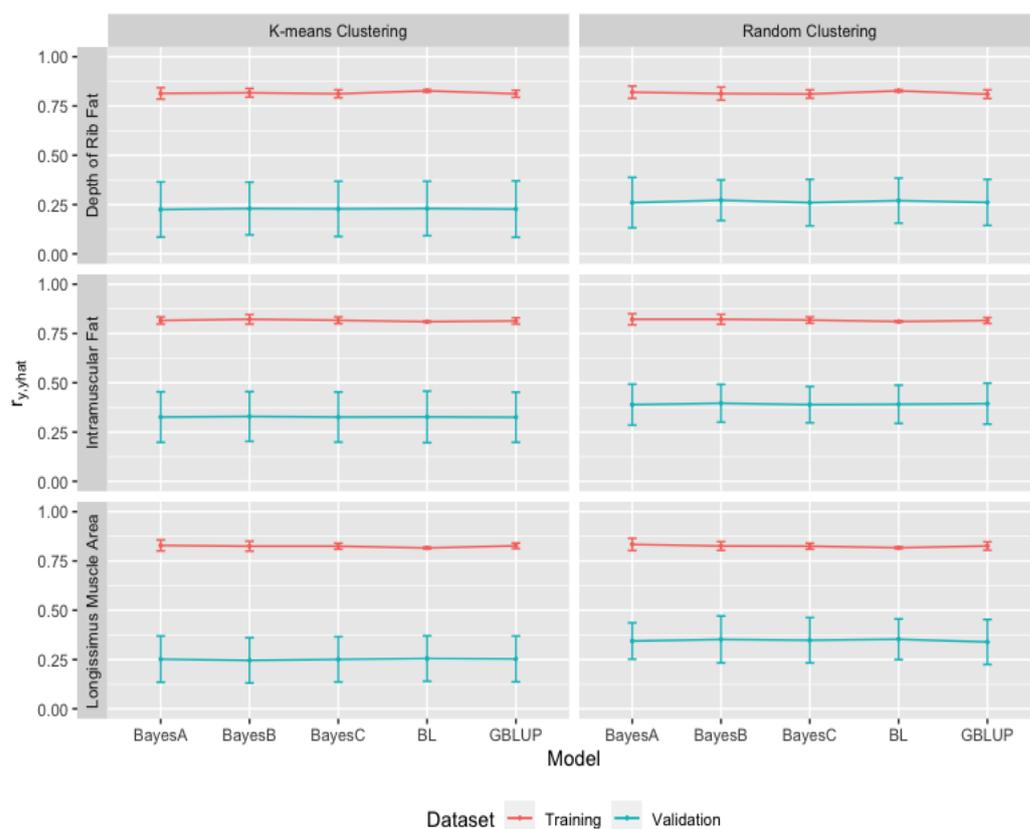
Figure 6. Predictive ability of Genomic Best Linear Unbiased Prediction (GBLUP) and the Bayesian (BayesA, BayesB, BayesC, and Lasso (BL)) methods for depth of rib fat (FAT), intramuscular fat (IMF), and longissimus muscle area (LMA) for carcass traits from 10-fold k-means and random cluster training datasets.

Figure 6 indicated that GBLUP (on average, 0.227 ± 0.143, 0.325 ± 0.127, 0.253 ± 0.116 and 0.261 ± 0.117, 0.394 ± 0.104, 0.339 ± 0.114), BayesA (on average, 0.225 ± 0.140, 0.326 ± 0.128, 0.252 ± 0.117 and 0.260 ± 0.128, 0.389 ± 0.104, 0.344 ± 0.092), BayesB (on average, 0.230 ± 0.134, 0.329 ± 0.126, 0.246 ± 0.115 and 0.272 ± 0.103, 0.396 ± 0.096, 0.352 ± 0.119), BayesC (on average, 0.228 ± 0.140, 0.326 ± 0.127, 0.251 ± 0.115 and 0.260 ± 0.118, 0.389 ± 0.092, 0.348 ± 0.115), and BL (on average, 0.230 ± 0.138, 0.327 ± 0.131, 0.255 ± 0.115 and 0.270 ± 0.114, 0.391 ± 0.097, 0.353 ± 0.103) methods within the 10-fold k-means and random cluster cross-validation datasets produced similar correlations for carcass (FAT, IMF and LMA) traits. The ranges of the correlations within 10-fold k-means and random cluster cross-validations were from 0.225 ± 0.140 to 0.230 ± 0.138 and 0.260 ± 0.128 to 0.272 ± 0.103 for FAT, 0.325 ± 0.127 to 0.329 ± 0.126 and 0.389 ± 0.104 to 0.396 ± 0.096 for IMF, and 0.246 ± 0.115 to 0.255 ± 0.115 and 0.339 ± 0.114 to 0.353 ± 0.103 for LMA. The ranges of correlations from the random cluster cross-validation were higher than those from the k-means cluster cross-validation. The trait of IMF produced higher correlations than the traits of FAT and LMA within carcass traits.

The predictive performances for growth (BW, WW, and YW) and carcass (FAT, IMF, and LMA) traits from k-means and random cluster training and validation datasets were found different within the GBLUP and the Bayesian methods, which depends on the genetic architecture of the traits. The similar predictive performances from the GBLUP and the Bayesian methods for growth (BW, WW and YW) and carcass (FAT, IMF and LMA) also suggested that the genetic structures of growth and carcass traits controlled by many genes with small effects. The carcass traits also resulted in the higher heritabilities and then higher predictive performances than growth traits. These results also revealed that the 10-fold k-means and random cluster cross-validation datasets resulted in significantly

lower correlations than training datasets for growth (BW, WW and YW) and carcass (FAT, IMF and LMA) traits within the GBLUP and the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods. The decrease in mean correlations for growth (BW, WW, and YW) and carcass (FAT, IMF, and LMA) traits ranged from 52% to 87% in the 10-fold k-means and random cluster cross-validation datasets. In addition, k-means cluster minimizing genetic relationship among cross-validation datasets produced lower correlations than the random cluster and the ranges of the decrease were from 16% to 22% for BW, 40% to 49% for WW, 23% to 26% for YW as growth traits, and 12% to 15% for FAT, 16% to 18% for IMF, and 25% to 30% for LMA as carcass traits across the GBLUP and the Bayesian methods.

The accuracies of GEBV from the GBLUP and the Bayesian methods were obtained from Pearson's correlations ($r_{y,\hat{y}}$) between observed and predicted phenotypes divided by the square root of the estimated heritabilities and given in Figure 7 for growth and in Figure 8 for carcass traits within k-means and random cluster cross-validation datasets. As seen in Figure 7 for growth traits, the accuracies from the GBLUP and the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods were 0.402 ± 0.033, 0.630 ± 0.033, 0.610 ± 0.033, 0.640 ± 0.033, 0.514 ± 0.034 within k-means clustering and 0.500 ± 0.042, 0.797 ± 0.031, 0.730 ± 0.043, 0.823 ± 0.029, 0.630 ± 0.036 within random clustering for BW; 0.220 ± 0.033, 0.425 ± 0.034, 0.386 ± 0.034, 0.416 ± 0.034, 0.288 ± 0.033 within k-means clustring and 0.406 ± 0.027, 0.702 ± 0.036, 0.386 ± 0.038, 0.768 ± 0.041, 0.589 ± 0.038 within random clustering for WW; and 0.447 ± 0.053, 0.633 ± 0.053, 0.620 ± 0.052, 0.645 ± 0.052, 0.618 ± 0.052 within k-means clustring and 0.600 ± 0.036, 0.820 ± 0.038, 0.870 ± 0.033, 0.844 ± 0.031, 0.817 ± 0.038 within random clustring for YW. As seen in Figure 8 for carcass traits, the accuracies from the GBLUP and the Bayesian (BayesA, BayesB, BayesC and Lasso) methods were 0.408 ± 0.045, 0.624 ± 0.044, 0.615 ± 0.042, 0.632 ± 0.044, 0.594 ± 0.044 within k-means clustering and 0.477 ± 0.037, 0.695 ± 0.040, 0.727 ± 0.033, 0.721 ± 0.037, 0.697 ± 0.036 within random clustering for FAT; 0.557 ± 0.040, 0.729 ± 0.040, 0.718 ± 0.040, 0.729 ± 0.040, 0.750 ± 0.041 within k-means clustering and 0.676 ± 0.033, 0.849 ± 0.033, 0.864 ± 0.030, 0.870 ± 0.029, 0.897 ± 0.031 within random clustering for IMF; and 0.428 ± 0.037, 0.594 ± 0.037, 0.580 ± 0.036, 0.592 ± 0.036, 0.618 ± 0.036 within k-means clustering and 0.573 ± 0.036, 0.789 ± 0.029, 0.830 ± 0.038, 0.820 ± 0.036, 0.856 ± 0.033 within random clustering for LMA.

The averaged accuracies of GEBV over all methods were 0.559 (0.696) for BW, 0.347 (0.645) for WW, 0.593 (0.790) for YW, 0.575 (0.663) for FAT, 0.697 (0.831) for IMF, and 0.562 (0.774) for LMA in k-means (random) cluster cross-validation datasets. As seen in Figures 7 and 8, the random clustering approach resulted in higher accuracies of GEBV (24% for BW, 87% for WW, 33% for YW, 15% for FAT, 19% for IMF, and 37% for LMA) than the k-means clustering approach because of the higher relationship between training and validation datasets in random clustering. Habier et al. [45] executed the genome-wise analysis of milk yield, fat yield, protein yield, and somatic cell score, and indicated that the accuracy of GEBV decreased by reducing the genomic relationship between animals for the training and validation datasets. Saatchi et al. [19] found the accuracies of 0.554 and 0.700 for BW, 0.333 and 0.534 for WW, 0.356 and 0.573 for YW, 0.603 and 0.793 for FAT, 0.690 and 0.817 for Marbling, and 0.601 and 0.694 for LMA in the k-means and random cross-validation datasets from Angus cattle, and they suggested that minimizing the genetic relationships between animals from training and validation sets using k-means clustering resulted in the conservative accuracies of GEBV.

Daetwyler et al. [46] determined that the high accuracy of GEBV resulted from the family relationships rather than LD between SNP and QTL in a multiple-breed sheep population. Chen et al. [47] showed that individuals with close relatives in the training population had a higher accuracy of GEBV. Kang et al. [48] also reported the decreasing accuracy of GEBV with an increasing generation gap between the training and validation datasets. Zhou et al. [49] studied the factors affecting GEBV accuracy and reported that the genetic relationship between animals from cross-validation datasets created a more

important effect than the LD between SNP and QTL on the accuracy of GEBV because the decrease in the accuracy of GEBV happened even when the LD between SNP increased.
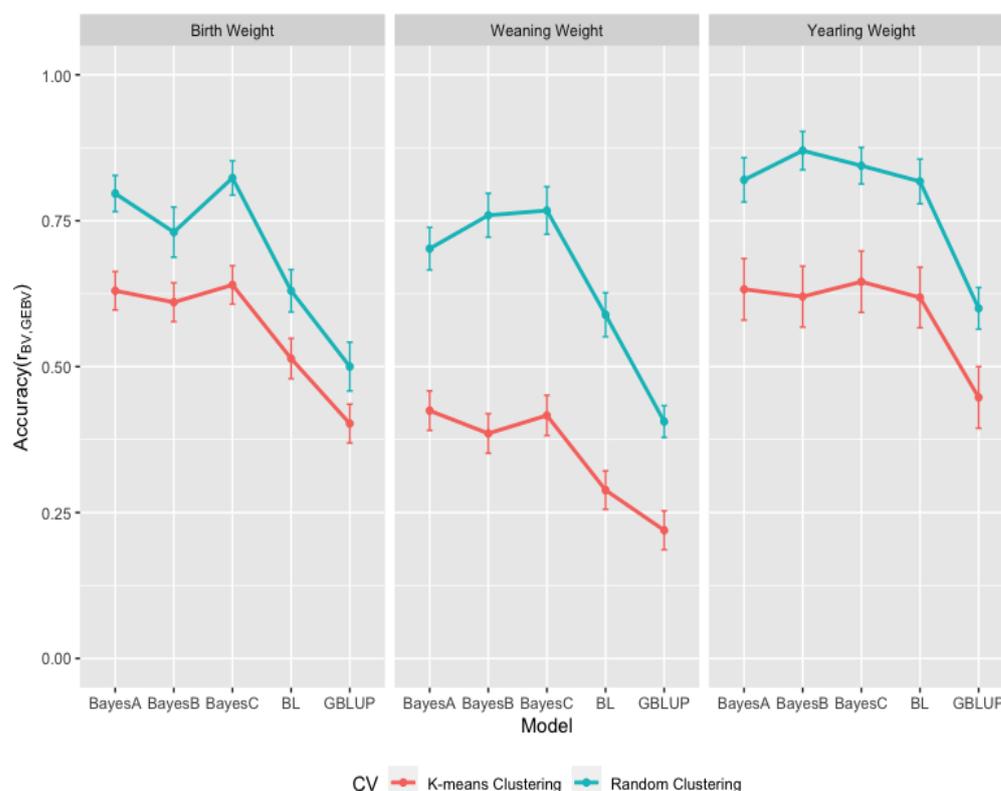
Figure 7. Accuracy of GEBV from Genomic Best Linear Unbiased Prediction (GBLUP) and the Bayesian (BayesA, BayesB, BayesC, and Lasso (BL)) methods for birth weight (BW), weaning weight (WW) and yearling weight (YW) for growth traits from 10-fold k-means and random cluster cross-validation datasets.

The accuracies of GEBV from the GBLUP and the Bayesian (BayesA, BayesB, BayesC, and Lasso) methods averaged over all k-means and random clustering were 0.451, 0.713, 0.670, 0.732, 0.572 for BW, 0.313, 0.563, 0.572, 0.592, 0.439 for WW, 0.524, 0.726, 0.745, 0.745, 0.718 for YW in the growth traits, 0.442, 0.659, 0.671, 0.677, 0.645 for FAT, 0.617, 0.789, 0.791, 0.799, 0.824 for IMF, and 0.500, 0.692, 0.705, 0.706, 0.737 for LMA in the carcass traits. As presented in Figures 7 and 8, the averaged accuracies of GEBV suggested that GBLUP resulted in lower accuracies of GEBV than the Bayesian methods within growth and carcass traits. The Bayesian methods exhibited quite similar accuracies of GEBV and the BayesC method for growth traits, and the Bayesian LASSO method for carcass traits provided higher accuracies of GEBV than other methods within the the Bayesian methods, respectively. Sun et al. [50] compared the GBLUP and the Bayesian methods using simulated data and found that the GBLUP had lower accuracy than BayesB and BayesCπ, and the Bayesian methods resulted in quite similar accuracies. Gao et al. [51] also reported that the Bayesian methods performed better accuracy of GEBV than the GBLUP methods in the genome analysis of milk production traits of Nordic Holstein cows. However, Chen et al. [52] reported that GBLUP performed better than Bayes B in the genomic analysis of carcass traits from Angus and Charolais beef cattle. Hayes et al. [53] also found that the GBLUP and the BayesB methods resulted in similar accuracies in a multibreed dairy population. Additionally, Ostersen et al. [54] reported no difference among the GBLUP, the Bayesian LASSO, and the Bayesian mixture methods based on 60,000 SNP data, and Ge et al. [55] reported the similar predictive accuracy for the GBLUP and the Bayesian methods for growth traits at weaning and yearling ages in Yaks. Although the predictive abilities of the GBLUP and the Bayesian methods were quite similar for growth and carcass traits and

k-means and random clusters (Figures 5 and 6) in the current study, the realized accuracies of the GEBV of the GBLUP and the Bayesian methods were not similar (Figures 7 and 8) because of the heritability estimates from the GBLUP and the Bayesian (BayesA, BayesB, BayesC and Lasso) methods. As described by Rolf et al. [42] for the smaller subsets of the data used in analyses, robust and reasonable heritability estimates can be obtained from GBLUP methodologies compared to the Bayesian methods. The accuracies of GEBV from the GBLUP in this study were then found in the range of the theoretical predicted accuracies between 0.26 and 0.34 based on the heritability estimates of traits from 0.25 to 0.40 [56].
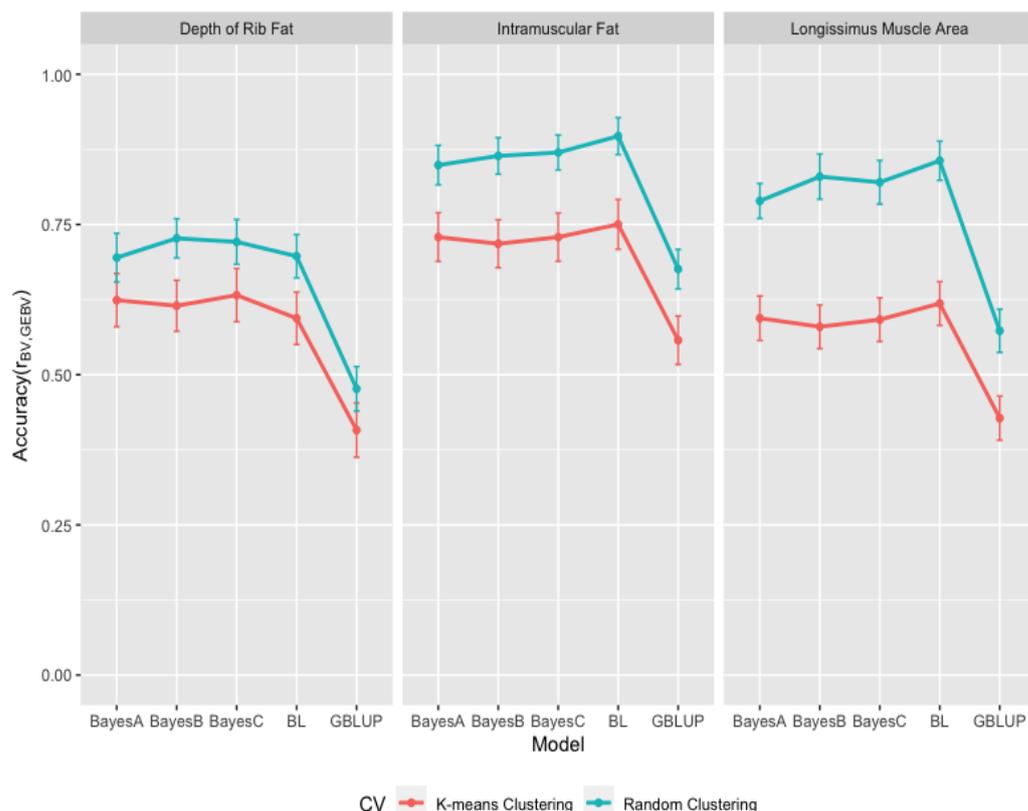
Figure 8. Accuracy of GEBV from Genomic Best Linear Unbiased Prediction (GBLUP) and the Bayesian (BayesA, BayesB, BayesC and Lasso (BL)) methods for depth of rib fat (FAT), intramuscular fat (IMF) and longissimus muscle area (LMA) for carcass traits from 10-fold k-means and random cluster cross-validation data sets.

4. Conclusions

In order to explore the translation of genomic prediction for growth and carcass traits in Brangus cattle, the GBLUP and the Bayesian (BayesA, BayesB, BayesC and Lasso) methods were used to estimate GEBV for growth (BW, WW and YW) and carcass (FAT, IMF and LMA) traits within k-means and random clusters in this study. The heritability estimates from k-means and random cluster were quite similar across genomic prediction methods. The Bayesian methods underestimated the heritabilities between 0.06 and 0.21; however, the heritability estimates from GBLUP were between 0.21 and 0.35 for growth and carcass traits, and they parallel these types of estimates in the literature. Including low-density SNP markers with low minor allele frequency would cause a poor performance to estimate the heritabilities of traits with the Bayesian methods compared with GBLUP using genomic relationship. K-means cluster appears to minimize the genetic relationships among cross-validation datasets and yields lower correlations than in a random cluster. These results of the current study suggested that the level of genetic relationship between the training and validation data influences the prediction ability of genomic selection

methods and the accuracy of GEBV. The prediction ability of the GBLUP and the Bayesian methods within k-means and random clusters were quite similar for growth and carcass traits; however, the Bayesian methods overestimated the accuracies of GEBV because of the lower estimates of the heritability of growth and carcass traits. However, the GBLUP resulted in more reasonable accuracy of GEBV for growth and carcass traits collected from Brangus heifers.

Author Contributions: Conceptualization, S.O.P., K.K., M.S. and M.G.T.; data curation, S.O.P. and K.K.; methodology, S.O.P., K.K., M.S. and B.M.; software, K.K., M.S. and B.M.; formal analysis, S.O.P., K.K., M.S., B.M. and A.K.T.; validation, K.K. and M.S.; writing—original draft preparation, S.O.P. and K.K.; writing—review and editing, S.O.P., K.K., M.S., B.M., A.K.T. and M.G.T. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the Institutional Animal Care and Use Committee Guidelines.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, [S.O.P.], upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **2001**, *157*, 1819–1829. [CrossRef]
2. Matukumalli, L.K.; Lawley, C.T.; Schnabel, R.D.; Taylor, J.F.; Allan, M.F.; Heaton, M.P.; O'Connell, J.; Moore, S.S.; Smith, T.P.L.; Sonstegard, T.S.; et al. Development and characterization of high density SNP genotyping assay for cattle. *PLoS ONE* **2009**, *4*, e5350. [CrossRef]
3. Applied Biosystems. Axiom Bovine Genotyping v3 Array (384HT format). 2019. Available online: https://www.thermofisher.com/order/catalog/product/55108%209#/551089 (accessed on 15 January 2022).
4. Illumina. Infinium iSelect Custom Genotyping Assays. 2016. Available online: https://www.illumina.com/content/dam/illumina-marketing/documents/products/technotes/technote_iselect_design.pdf (accessed on 15 January 2022).
5. Habier, D.; Fernando, R.L.; Kizilkaya, K.; Garrick, D.J. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinform.* **2011**, *12*, 186. [CrossRef]
6. Pérez, P.; de los Campos, G. Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **2014**, *198*, 483–495. [CrossRef]
7. Hayes, B.; Bowman, P.; Chamberlain, A.; Goddard, M. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* **2009**, *92*, 433–443. [CrossRef] [PubMed]
8. VanRaden, P.M.; Van Tassell, C.P.; Wiggans, G.; Sonstegard, T.; Schnabel, R.; Taylor, J.F.; Schenkel, F.S. Invited review: Reliability of genomic predictions for north American Holstein bulls. *J. Dairy Sci.* **2009**, *92*, 16–24. [CrossRef]
9. Peters, S.O.; Kizilkaya, K.; Garrick, D.J.; Fernando, R.L.; Reecy, J.M.; Weaber, R.L.; Silver, G.A.; Thomas, M.G. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. *J. Anim. Sci.* **2012**, *90*, 3398–3409. [CrossRef] [PubMed]
10. Peters, S.O.; Kizilkaya, K.; Garrick, D.J.; Fernando, R.L.; Reecy, J.M.; Weaber, R.L.; Silver, G.A.; Thomas, M.G. Heritability and Bayesian genome-wide association of binary traits of first service conception and heifer pregnancy in Brangus heifers. *J. Anim. Sci.* **2013**, *91*, 605–612. [CrossRef]
11. de Los Campos, G.; Hickey, J.M.; Pong-Wong, R.; Daetwyler, H.D.; Calus, M.P.L. Whole genome regression and prediction methods applied to plant and animal breeding. *Genetics* **2013**, *193*, 327–345. [CrossRef]
12. Meuwissen, T.; Hayes, B.; Goddard, M.E. Accelerating improvement of livestock with genomic selection. *Annu. Rev. Anim. Biosci.* **2013**, *1*, 221–237. [CrossRef] [PubMed]

13. Meuwissen, T.; Hayes, B.; Goddard, M.E. Genomic selec-tion: A paradigm shift in animal breeding. *Anim. Front.* **2016**, *6*, 6–14. [CrossRef]

14. Crossa, J.; de Los Campos, G.; Pérez, P.; Gianola, D.; Burgueño, J.; Araus, J.L.; Makumbi, D.; Singh, R.P.; Dreisigacker, S.; Yan, J.; et al. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* **2010**, *186*, 713–724. [CrossRef] [PubMed]

15. Colombani, C.; Legarra, A.; Fritz, S.; Guillaume, F.; Croiseau, P.; Ducrocq, V.; Robert-Granié, C. Application of bayesian least absolute shrinkage and selection operator (LASSO) and BayesCπ methods for genomic selection in French holstein and montbéliarde breeds. *J. Dairy Sci.* **2013**, *96*, 575–591. [CrossRef]

16. Esfandyari, H.; Sørensen, A.; Bijma, P.A. Crossbred reference population can improve the response to genomic selection for crossbred performance. *Genet. Sel. Evol.* **2015**, *47*, 76. [CrossRef] [PubMed]

17. Liu, H.; Zhou, H.; Wu, Y.; Li, X.; Zhao, J.; Zuo, T.; Zhang, X.; Zhang, Y.; Liu, S.; Shen, Y.; et al. The impact of genetic relationship and linkage disequilibrium on genomic selection. *PLoS ONE* **2015**, *10*, e0132379. [CrossRef] [PubMed]

18. Habier, D.; Fernando, R.L.; Dekkers, J.C.M. The impact of genetic relationship information of genome-assisted breeding values. *Genetics* **2007**, *177*, 2389–2397. [CrossRef]

19. Saatchi, M.; McClure, M.C.; McKay, S.D.; Rolf, M.M.; Kim, J.; Decker, J.E.; Taxis, T.M.; Chapple, R.H.; Ramey, H.R.; Northcutt, S.L.; et al. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. *Genet. Sel. Evol.* **2011**, *43*, 40. [CrossRef]

20. Villumsen, T.; Janss, L.; Lund, M. The importance of haplotype length and heritability using genomic selection in dairy cattle. *J. Anim. Breed Genet Z. Tierz. Zucht.* **2009**, *126*, 3–13. [CrossRef]

21. Clark, S.; Hickey, J.; van der Werf, J. Different models of genetic variation and their effect on genomic evaluation. *Genet Sel. Evol.* **2011**, *43*, 18. [CrossRef]

22. Luna-Nevarez, P.; Bailey, D.W.; Bailey, C.C.; VanLeeuwen, D.M.; Enns, R.M.; Silver, G.A.; DeAtley, K.L.; Thomas, M.G. Growth characteristics, reproductive performance, and evaluation of their associative relationships in Brangus cattle managed in a Chihuahuan Desert production system. *J. Anim. Sci.* **2010**, *88*, 1891–1904. [CrossRef]

23. Fortes, M.R.S.; Snelling, W.M.; Reverter, A.; Nagaraji, S.H.; Lehnert, S.A.; Hawken, R.J.; DeAtley, K.L.; Peters, S.O.; Silver, G.A.; Rincon, G.; et al. Gene network analyses of first service conception in Brangus heifers: Use of genome and trait associations, hypo- thalamic-transcriptome information, and transcription factors. *J. Anim. Sci.* **2012**, *90*, 2894–2906. [CrossRef] [PubMed]

24. Granato, I.S.C.; Galli, G.; de Oliveira Couto, E.G.; e Souza, M.B.; Mendonça, L.F.; Fritsche-Neto, R. snpReady: A tool to assist breeders in genomic analysis. *Mol. Breeding* **2018**, *38*, 102. [CrossRef]

25. Yin, L.; Zhang, H.; Tang, Z.; Xu, J.; Yin, D.; Zhang, Z.; Yuan, X.; Zhu, M.; Zhao, S.; Li, X.; et al. rMVP: A Memory-efficient, Visualization-enhanced, and Parallel-accelerated tool for Genome-Wide Association Study. *Genom. Proteom. Bioinform.* **2021**, *19*, 619–628. [CrossRef] [PubMed]

26. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: https://www.R-project.org/ (accessed on 10 March 2022).

27. Wimmer, V.; Albrecht, T.; Auinger, H.J.; Schon, C.C. Synbreed: A framework for the analysis of genomic prediction data using R. *Bioinformatics* **2012**, *28*, 2086–2087. [CrossRef]

28. vanRaden, P.M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **2008**, *91*, 4414–4423. [CrossRef]

29. Park, T.; Casella, G. The bayesian lasso. *J. Am. Stat. Assoc.* **2008**, *103*, 681–686. [CrossRef]

30. Henderson, C.R. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* **1976**, *32*, 69. [CrossRef]

31. Gorjanc, G.; Henderson, D.A.; Kinghorn, B.; Percy, A. GeneticsPed: Pedigree and Genetic Relationship Functions. 2020. R Package Version 1.52.0. Available online: http://rgenetics.org (accessed on 10 March 2022).

32. Kassambara, A.; Mundt, F. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. 2017. R Package Version 1.0.5. Available online: https://CRAN.R-project.org/package=factoextra (accessed on 10 March 2022).

33. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *Appl. Stat.* **1979**, *28*, 100–108. [CrossRef]

34. Legarra, A.; Robert-Granié, C.; Manfredi, E.; Elsen, J.M. Performance of genomic selection in mice. *Genetics* **2008**, *180*, 611–618. [CrossRef]

35. Mustafa, H.; Ahmad, N.; Heather, H.J.; Eui-soo, K.; Khan, W.A.; Ajmal, A.; Javed, K.; Pasha, T.N.; Ali, A.; Kim, J.J.; et al. Whole genome study of linkage disequilibrium in Sahiwal cattle. *S. Afr. J. Anim. Sci.* **2018**, *48*, 353–360. [CrossRef]

36. El Hou, A.; Rocha, D.; Venot, E. Long-range linkage disequilibrium in French beef cattle breeds. *Genet. Sel. Evol.* **2021**, *53*, 1–14. [CrossRef] [PubMed]

37. Singh, A.; Kumar, A.; Mehrotra, A.; Pandey, A.K.; Mishra, B.P.; Dutt, T. Estimation of linkage disequilibrium levels and allele frequency distribution in crossbred Vrindavani cattle using 50K SNP data. *PLoS ONE* **2021**, *16*, 1–10. [CrossRef] [PubMed]

38. Lu, D.; Sargolzaei, M.; Kelly, M.; Li, C.; Vander Voort, G.; Wang, Z.; Plastow, G.; Moore, S.; Miller, S.P. Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Front. Gene.* **2012**, *3*, 1–10. [CrossRef]

39. McKay, S.D.; Schnabel, R.D.; Murdoch, B.M.; Matukumalli, L.K.; Aerts, J.; Coppieters, W.; Crews, D.; Neto, E.D.; Gill, C.A.; Gao, C.; et al. Whole genome linkage disequilibrium maps in cattle. *BMC Genet.* **2007**, *8*, 74. [CrossRef]

40. Ríos-Utrera, Á.; Vega-Murillo, V.E.; Martínez-Velázquez, G.; Montaño-Bermúdez, M. Comparison of models for the estimation of variance components for growth traits of registered limousin cattle. *Trop. Subtrop. Agroecosyt.* **2011**, *14*, 667–674.

41. Neser, F.W.C.; van Wyk, J.B.; Fair, M.D.; Lubout, P.; Crook, B.J. Estimation of genetic parameters for growth traits in Brangus cattle. *S. Afr. J. Anim. Sci.* **2012**, *42*, 469–473. [CrossRef]

42. Rolf, M.M.; Garrick, D.J.; Fountain, T.; Ramey, H.R.; Weaber, R.L.; Decker, J.E.; Pollak, E.J.; Schnabel, R.D.; Taylor, J.F. Comparison of Bayesian models to estimate direct genomic values in multi-breed commercial beef cattle. *Genet. Sel. Evol.* **2015**, *47*, 1–14. [CrossRef] [PubMed]

43. Pires, B.C.; Tholon, P.; Buzanskas, M.E.; Sbardella, A.P.; Rosa, J.O.; Campos da Silva, L.O.; Torres Júnior, R.A.A.; Munari, D.P.; Alencar, M.M. Genetic analyses on bodyweight, reproductive, and carcass traits in composite beef cattle. *Anim. Prod. Sci.* **2017**, *57*, 415–421. [CrossRef]

44. Boldt, R.J. *Genetic Parameters for Fertility and Production Traits in Red Angus Cattle*; Master of Science, Colorado State University: Fort Collins, CO, USA, 2017.

45. Habier, D.; Tetens, J.; Seefried, F.R.; Lichtner, P.; Thaller, G. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genet. Sel. Evol.* **2010**, *42*, 5. [CrossRef]

46. Daetwyler, H.D.; Kemper, K.E.; Jh, V.D.W.; Hayes, B.J. Components of the accuracy of genomic prediction in a multi-breed sheep population. *J. Anim. Sci.* **2012**, *90*, 3375–3384. [CrossRef]

47. Chen, L.; Li, C.; Sargolzaei, M.; Schenkel, F. Impact of genotype imputation on the performance of GBLUP and Bayesian methods for genomic prediction. *PLoS ONE* **2014**, *9*, e101544. [CrossRef]

48. Kang, H.; Zhou, L.; Mrode, R.; Zhang, Q.; Liu, J.F. Incorporating single-step strategy into random regression model to enhance genomic prediction of longitudinal trait. *Heredity* **2016**, *119*, 459–467. [CrossRef]

49. Zhou, L.; Mrode, R.; Zhang, S.; Zhang, Q.; Li, B.; Liu, J. Factors affecting GEBV accuracy with single-step Bayesian models. *Heredity* **2018**, *120*, 100–109. [CrossRef] [PubMed]

50. Sun, X.; Habier, D.; Fernando, R.L.; Garrick, D.J.; Dekkers, J.C. Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. *BMC Proc.* **2011**, *5*, S13. [CrossRef]

51. Gao, H.; Su, G.; Janss, L.; Zhang, Y.; Lund, M. Model comparison on genomic predictions using high-density markers for different groups of bulls in the Nordic Holstein population. *J. Dairy Sci.* **2013**, *96*, 4678–4687. [CrossRef]

52. Chen, L.; Vinsky, M.; Li, C. Accuracy of predicting genomic breeding values for carcass merit traits in Angus and Charolais beef cattle. *Anim. Genet.* **2014**, *46*, 55–59. [CrossRef]

53. Hayes, B.J.; Bowman, P.J.; Chamberlain, A.C.; Verbyla, K.; Goddard, M.E. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* **2009**, *41*, 51. [CrossRef]

54. Ostersen, T.; Christensen, O.F.; Henryon, M.; Nielsen, B.; Su, G.; Madsen, P. Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in purebred pigs. *Genet. Sel. Evol.* **2011**, *43*, 38. [CrossRef]

55. Ge, F.; Jia, C.; Bao, P.; Wu, X.; Liang, C.; Yan, P. Accuracies of Genomic Prediction for Growth Traits at Weaning and Yearling Ages in Yak. *Animals* **2020**, *10*, 1793. [CrossRef] [PubMed]

56. Pryce, J.E.; Arias, J.; Bowman, P.J.; Davis, S.R.; Macdonald, K.A.; Waghorn, G.C.; Wales, W.J.; Williams, Y.J.; Spelman, R.J.; Hayes, B.J. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. *J. Dairy Sci.* **2011**, *95*, 2108–2119. [CrossRef] [PubMed]